

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2002, 23:33:11 ; Search time 1527.5 seconds  
(without alignments)  
16168.980 Million cell updates/sec

Title: US-10-025-514-15  
Perfect score: 1525  
Sequence: 1 tctagaccatggaagaccct.....ccagtaaggcctagtgcac 1525

Scoring table: IDENTITY\_NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_estc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est4:\*

13: gb\_est5:\*

14: gb\_est6:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pin:\*

21: em\_gss\_vit:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_oth:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	429	28.1	2478	11 AF130068	AF130068 Homo sapi
2	427.4	28.0	2571	11 AF113676	AF113676 Homo sapi
3	331.2	21.7	1392	11 AK002537	AK002537 Mus muscu
4	318.4	20.9	1296	11 AK004999	AK004999 Mus muscu
5	309.8	20.3	887	11 BQ643710	BQ643710 AGENCOURT
6	305.4	20.0	907	14 BQ648909	BQ648909 AGENCOURT

7	286.8	18.8	895	14	BQ653587	AGENCOURT
8	285.2	18.7	982	14	BQ646142	AGENCOURT
9	284.2	18.6	985	14	BM924019	AGENCOURT
10	283.8	18.6	924	14	BQ958958	AGENCOURT
11	274.8	18.0	1194	14	BM924813	AGENCOURT
12	273.4	17.9	907	14	BQ650189	AGENCOURT
13	270.4	17.7	959	14	BQ648524	AGENCOURT
14	268.6	17.6	891	14	BQ646548	AGENCOURT
15	268	17.6	796	12	BG567260	AGENCOURT
16	266.4	17.5	833	14	BQ007663	AGENCOURT
17	264.8	17.4	991	14	BQ064738	AGENCOURT
18	263.6	17.3	687	10	AV649258	AV649258
19	262	17.2	686	12	BG569467	AGENCOURT
20	261	17.1	917	14	BQ650502	AGENCOURT
21	260.2	17.1	916	14	BQ644130	AGENCOURT
22	260.2	17.1	977	14	BQ654015	AGENCOURT
23	259.6	17.0	740	10	AV653778	AV653778
24	259.6	17.0	944	14	BQ645533	AGENCOURT
25	259.6	17.0	1012	14	BQ064473	AGENCOURT
26	256	16.8	869	13	BI219058	AGENCOURT
27	253	16.5	654	12	BG563537	AGENCOURT
28	251.4	16.5	884	12	BG563548	AGENCOURT
29	251.2	16.5	917	14	BQ650698	AGENCOURT
30	250.6	16.4	746	12	BG566873	AGENCOURT
31	250.2	16.4	745	9	AI114643	AI114643
32	249.8	16.4	1044	9	AI528087	AI528087
33	249.6	16.4	813	13	BI759000	BI759000
34	249.4	16.4	764	12	BF384318	BF384318
35	249	16.3	731	10	AV649144	AV649144
36	249	16.3	813	14	BQ182052	AGENCOURT
37	246	16.1	804	12	BG618951	AGENCOURT
38	244.4	16.0	780	9	AI110706	AI110706
39	244.2	16.0	992	13	BI246204	AGENCOURT
40	243.8	16.0	752	12	BG618001	AGENCOURT
41	243.2	15.9	674	14	BM724546	BM724546
42	243.2	15.9	759	12	BG533293	AGENCOURT
43	243	15.9	749	9	AI526727	AI526727
44	242	15.9	629	12	BG545257	AGENCOURT
45	241.2	15.8	780	12	BG617594	AGENCOURT

## ALIGNMENTS

RESULT 1	AF130068	Homo sapiens clone FLB8226	2478 bp	mrna	linear	HTC 08-MAY-2001
LOCUS	AF130068	Homo sapiens clone FLB8226	2478 bp	mrna	linear	HTC 08-MAY-2001
DEFINITION	AF130068	Homo sapiens clone FLB8226	2478 bp	mrna	linear	HTC 08-MAY-2001
ACCESSION	AF130068	Homo sapiens clone FLB8226	2478 bp	mrna	linear	HTC 08-MAY-2001
VERSION	AF130068.1	GI:11493442				
KEYWORDS	HTC					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 2478)					
AUTHORS	Zhang, C., Yu, Y., Zhang, S., Wei, H., Bi, J., Zhou, G., Dong, C., Zai, Y., Xu, W., Gao, F., Liu, M. and He, F.					
TITLE	Functional prediction of the coding sequences of 75 new genes deduced by analysis of cDNA clones from human fetal liver					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 2478)					
AUTHORS	Zhang, C., Yu, Y., Zhang, S., Wei, H., Bi, J., Zhou, G., Dong, C., Zai, Y., Xu, W., Gao, F., Liu, M. and He, F.					
TITLE	Direct Submission					
JOURNAL	Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing, Beijing 100850, P. R. China					
FEATURES	Location/Qualifiers					
source	1..2478					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/clone="FLB8226"					

```

/tissue_type="liver"
/dev_stage="fetus"
1082..2338
/notes="predicted protein of H02209"
/codon_start=1
/evidence-not_experimental
/product="PRO2209"
/protein_id="A035496.1"
/db_xref="GI:11493443"
/translation="MPSVSWGILLIAGLCCLVPVSLAEDPOGDAAQAKTDTSHHDDH
PTNKITPNIAEAFSLYQIAHQSNSTNIFSPVSIATAFAFAMLSIGTKADTHDELE
GLNFNTEPEAGIHEGFOELLRLNQPSQLQTLTNGFLFSEGLKLVDFKLEVDK
LYSEATFVNGEERAKKOINDYVEKGTQKIVDLVKELDRDTFALVNYIFFKGGK
ERFEVNDTEEDPHVDQYTVKPMKRLGMENIOHCKLSSWLLMYLGNATAIF
FLPDGKLQHLNELTHDIITKFELENDRSLHLPKLSITGYDLKSVLGOLGTTK
VFSNGADLSVTEAPLKLSKAVHKAVLTIDEAGTEAAGAMFLEALPMSIPPEVKNK
PFVFLMEQNTKSPFMGKVVNPTQK"
BASE COUNT 650 a 627 c 597 g 604 t
ORIGIN

Query Match 28.18; Score 429; DB 11; Length 2478;
Best Local Similarity 60.28; Pred. No. 6.6e-101; Indels 0; Gaps 0;
Matches 711; Conservative 0; Mismatches 470;

QY 12 GAAGACCTCAAGGCGAGCGCGCTCAAAAAACCGACACAGTCATCAGACCAAGACAT 71
DB 114 GAGGATCCCAGGAGATGCTGCCAGAGACAGATACATCCACCATGATCAGGATCAC 1213
QY 72 CCGACTTTTAAATAATTACTCCAAATTTAGCCGNAATTTGCTTTTCTTCTATAGACAA 131
DB 1214 CCAACCTTCAACAAGATCACCCCAACCTGGCTGAGTTGCGCTTCAGCCTATACGCCAG 1273
QY 132 TTAGTCCATCAAGTAATTTCTACTAACATTTTTTTTACTGCTGTTCTTATTTGCCACTGCT 191
DB 1274 CTGGACACCATCCACAGCACCATAATCTCTTCTCCCGAGTGAGCATCGCTACAGCC 1333
QY 192 TTGCGCATGTTGAGTTTGGTACTAAAGCCGATACCCATGAGGAGATTTTGAAGGTTTA 251
DB 1334 TTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCAGATGAAATCTGGAGGGCGCTG 1393
QY 252 AACTTTAATTGACCAATCCAGAACCCCAATTCAGAGGGTTTTCAGAGTTGTTG 311
DB 1394 AATTTCACCTCAGGAGATTCGGAGGCTCAGATCCATGAGGCTTCCAGGAATCTCTC 1453
QY 312 AGAATTTTGAATCAACCTGATTTCTCAATTTGCAATTAATCTACTGTTAAGCGTTTATTTG 371
DB 1454 CGTACCTCACCAGCCAGACAGCCAGCTCCAGCTGACCCGCAATGGCCTGTTCTCTC 1513
QY 372 TCTGAAGGTTTAAATTTGGTTGACAAATTCCTAGAGAGCGTCAAGAACTATATCATAGT 431
DB 1514 AGCGAGGGCTGAAGCTAGTGGATAAGTTTGGAGGATGTTAAAGATTTGACCACTCA 1573
QY 432 GAGGCTTTTACCTTAATTTGTGTGATCTAGGAGAGCTTAAAGCAAAATTAATGATTAT 491
DB 1574 GAAGCCTTCACTGTCAACTTCGGGGACACCGAGAGGCCCAAGAACAGATCAACGATTAC 1633
QY 492 GTTGAGAAAGGCCACCGAGGTAAGATCGTTGACCTAGTTTAAAGAAATAGATCGTGATAGT 551
DB 1634 GTGAGAGAGGCTACTCAAGGGAATAATTTGTGATTTTGGTCAAGGAGCTTGACAGACACA 1693
QY 552 GTCTTCGCATAGTTTAATATATTTTTTCAAGGGTAAAGTGGGAAGCTCTTTCGAGGTT 611
DB 1694 GTTTTTCTGTGTGATTTACATCTTCTTTAAAGGCAATTTGGAGAGACCCCTTTGAAGTC 1753
QY 612 AAGATATCTGAAGGAGAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAAATG 671
DB 1754 AAGGACACCGAGGAAGAGGACTTCCAGCTGGACCGAGTGCACCGTGAAGGTGCCTATG 1813
QY 672 ATGAAAGACTGGGTATGTCAATATTTCAATATTTCAAAAAATAAGTTCTTGGGCTTA 731
DB 1814 ATGAGCGTTTGAAGCATGTTTAACTCCAGCACTGTGAAGAGCTGTCCAGCTGGGCTG 1873
QY 732 TTAATGAAGTATTTAGGTAAAGCTACTGCTATTTTTTTTTTACCAGACGAAAGTAAAGCTT 791

```

```

DB 1874 CTGATGAATACCTGGGCAATGCCACCGCCATCTTCTTCTCGCTGATGAGGGAACTA 1933
QY 792 CAACATTTAGAGAATGAGTTGACTCATGACATTAATTAATTTTATAGAACGAGAT 851
DB 1934 CAGCAGCTGGAATGAATCACTCCACGATATCATCACCAGTTCTCTGGAATGAAGAC 1993
QY 852 CGTCGTAGCGCTTCTCTGCACTGCCCAAGTTAAGTATCACCGGTACTTACGACTTAAA 911
DB 1994 AGAAGCTCTGCCAGCTTACATTTACCAAACTGTCCATTTACTGGAACCTATATCTGAAG 2053
QY 912 TCTGTTTTAGGCCAGTTAGTATTAACCAAGTTTTTTTCTAACGGTCCGCTATTTGAGTGT 971
DB 2054 AGCTCCTGGGTCAACTGGGCACTACATAAGTCTTTCAGCAATGGGCTGACCTCTCCGGG 2113
QY 972 GTTACTGGAAGAAGCTCCATTAATAATTTAGTAAGTGTTCACAAAGCCGCTTCTTAATTA 1031
DB 2114 GTCACAGAGGAGCCACCCCTGAGCTCTCAAGGCGGTGCATAAGGCTGTGCTGACCATC 2173
QY 1032 GATGAAAAGGTTACGAGGCGCGCGCTATGTTCTCTGGAAGCTATTTCCAATGAGCAT 1091
DB 2174 GAGGAAAAGGACTGAAGCTGCTGGGCGCATGTTTTTAGAGGCCATACCATGCTATC 2233
QY 1092 CCACGAGAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATGAGCAGACACTAAA 1151
DB 2234 CCCCCGAGTCAAGTTCAACAAACCTTTGCTCTTCTTAATGATTGAACAAAATACCAAG 2293
QY 1152 AGCCCATGTTTATGTTAGGTAAGTTGTCAACCCCAACTCAGAA 1192
DB 2294 TCTCCCTCTTCTGAGAAAGTGGTGAATCCCAACCCAAAA 2334

RESULT 2
AF113676 2571 bp mRNA linear HTC 08-MAY-2001
LOCUS Homo sapiens clone FLB2803 PRO0684 mRNA, complete cds.
DEFINITION AF113676
ACCESSION AF113676
VERSION AF113676.1 GI:6855600
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2571)
AUTHORS Zhang, C., Yu, Y., Zhang, S., Ouyang, S., Luo, L., Wei, H., Zhou, G., Zhou, W., Bi, J., Zhang, Y., Liu, M. and He, F.
TITLE Functional prediction of the coding sequences of 32 new genes deduced by analysis of cDNA clones from human fetal liver
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2571)
AUTHORS Zhang, C., Yu, Y., Zhang, S., Ouyang, S., Luo, L., Wei, H., Zhou, G., Zhou, W., Bi, J., Zhang, Y., Liu, M. and He, F.
TITLE Direct Submission
JOURNAL Submitted (14-DEC-1998) Department of Experimental Hematology, Institute of Radiation Medicine, Beijing Taping Road 27, Beijing 100850, P. R. China
FEATURES
source
1. 2571
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FLB2803"
/tissue_type="liver"
/dev_stage="fetus"
/notes="H00684"
515..819
/rpc_family="Alu"
repeat_region
1109..2365
/notes="predicted protein of H0684"
CDS
/codon_start=1
/evidence-not_experimental
/product="PRO0684"
/protein_id="AAF29581.1"
/db_xref="GI:6855601"

```

BASE COUNT  
ORIGIN  
800 a 556 g 577 t  
638 c  
PFFGLMEQNTKSPFMGVNPTQK"  
/translation="MSSVSVSGILLAGCLLPVSLAEDPQGDAQKTDTSHHDDQH  
PTFNKIPNLAFAFSPYQLARQSNSTIFFSPVSIATAFAMLSLQTKAADHTDEILE  
GNLNFITPEAQIHFGFOLRLTNPDQSLOLITGNGLESLGLVKDFLEDEYVK  
LYHSEATVNFNGDEEAKQINDYKEKQTQKIVDLKELDRDTFVALVNIFFKGK  
ERPEVKDTEEDFHDVQTVVPMKMRGMFNHQCKLSSWMLMKYLGNLATAIF  
FLPDEGKLHLENELHTIITFLNEORRSASHLPKLSITGYTDKSLVGLQGTIT  
VFNGLADLCTEAEAKLSKAVHKAVLTIDEKGTAAAGAMFLEAIPMSIPPEVKENK





\_\_\_\_\_

\_\_\_\_\_



using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT	242 a	224 c	232 g	187 t	2 others
ORIGIN					
Query Match	20.3%; Score 309.8; DB 14; Length 887;				
Best Local Similarity	60.0%; Pred. No. 6.5e-70;				
Matches	532; Conservative 0; Mismatches 354; Indels 1; Gaps 1;				
QY	195	GCATGTTGAGTTT	AGGTACTAAAGCCGATACCCATGACGAGATTTT	TAGAAGTTTAAAC	254
Db	1	GCAATGCTCTCCCT	GGGACCAAGGCTGACACTCAGATGAATCTCTGGAGGCTGAAT	60	
QY	255	TTTAATTTACCGAATCC	CAGAACCCAAATTCACGAGGGTTTTC	CAAGAGTTGTTGAGA	314
Db	61	TTCAACCTCACGGAGATT	CCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACCTCCCGT	120	
QY	315	ACTTTGAATCAACCTGATTCT	CAATTTGCAATTTAACTACTGTTAAGCGTTTATTTTGTCT	374	
Db	121	ACCTCAACCGCAGACAGCC	AGCTCCAGCTGACCAACCGCAATGGCCTGTTCTCAGC	180	
QY	375	GAAAGTTTAAATTTGGT	GACAAATTCCTAGAACGCTCAAGAACTATATCATAGTGAG	434	
Db	181	GAGGCGCTGAAGCTAGT	AGTAAAGTTTGGAGGATGTTAAAAAGTTTGTACCACTCAGAA	240	
QY	435	GCCTTTTACCGTTAAATTT	GGTGATACTGAGGAAGCTAAAGCAAAATTAAGTATATGTT	494	
Db	241	GCCTTCACTGTCACTT	CGGGACACCAAGAGGCCAAGAACAGATCAACGATTTACGTG	300	
QY	495	GAGAAAGCCACCCAGGGT	AAAGTCTTGACCTAGTTAAAGAAATTAGATCGTGATACCGTC	554	
Db	301	GAGAAAGGTACTCAAGG	GAATTTGCTGATGCTGATGCTGATGCTGATGCTGATGCTG	360	
QY	555	TTCCGCACTAGTTAACT	ATATTTTTCAGAGGTAGTGGGAACGCTCTTCCAGGTTAAA	614	
Db	361	TTTGTCTGGTGAAT	TACATCTCTTTTAAAGCAATGGGAGACACCTTTTGAAGTCAAG	420	
QY	615	GATACTGAAGAGAGATT	TTTCATCTTCATCAAGTTACTACTGTCAAAAGTTCCATGATG	674	
Db	421	GACACCGAGGAGAGACT	TTCCAGCTGGACCGGTGACCCCGTGAAGGTCCTATGATG	480	
QY	675	AAAGACTGGGTATGTTCA	ATATTAACATTCACCAATTCGAAATAATTAAGTTCTTTGGTCTTATTA	734	
Db	481	AAGCGTTTAGGCATCTT	TAACATCCAGCACTGTAAGAGCTGTCAGCTGGTCTGCTG	540	
QY	735	ATGAAGTATTTAGGTAA	CGCTACTGCTATTTTATTTTACAGAGGAAGGTAAGCTTCAA	794	
Db	541	ATGAATACTCGGGCAAT	GGCCACCGCCATCTTCTTCTGCTGATGAGGGGAAACACAG	600	
QY	795	CATTTAGAGAATGAGT	TGACTCATGACATATTTACTAAATTTTATAGAACGAGGATCGT	854	
Db	601	CACCTGGAAATGACT	CACCCACGATATCATCAAGTTCTCTGGAAATGAAGACAGA	660	
QY	855	CGTAGCGTCTCTGCACT	CCCAAGTTAAGTATCACCGGTACTTACGACTTAAATCT	914	
Db	661	AGGTCTGCCAGTACAT	TTTACCACAACTGTCCATTACTGGAACCTATGATCTGAAGC	720	
QY	915	GTTTTAGCCAGTAGT	ATTTACCAAGTTTCTTCTAACGGTCCGATTTGAGTGGTGT	974	
Db	721	GTCTTGTGCTAACCT	AGGTCATCAAGGTTTTCAGCAATGGGCTGACCTCTCCGGNGTC	780	
QY	975	ACTGAAGAGCTCCAT	TAAATTTAGTAAAGCTGTTCAAGCGGCTTTAACTATTGAT	1034	
Db	781	ACAGAGGAGGACCCCT	GAACCTCTCCAGGC - GTGCATAAGGCTGTGCTGACCATCGAC	839	
QY	1035	GAAAGGTTACCGAGG	CGCGCGGCTATGTTCTCTGGAAGCTATTCC	1081	
Db	840	GAGAAAGGCTGAAGCT	GTGCTGGGCCATGTTTTAGAGGCCATACC	886	
RESULT 6					
BQ648909					

LOCUS  
DEFINITION  
BO648909  
AGENCOURT\_8351269 NIH\_MGC\_100 Homo sapiens cDNA clone IMAGE:6286839  
5', mRNA sequence.  
BO648909  
BO648909.1 GI:21773081  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 907)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: CGAP (Stanford)  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1CM2488 row: e column: 16  
High quality sequence stop: 721.

FEATURES  
source  
1..907  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6286839"  
/clone\_lib="NIH\_MGC\_100"  
/tissue\_type="hepatocellular carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: liver; Vector: pOTB7; Site:1: XhoI; Site:2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GSCACAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

BASE COUNT 246 a 228 c 239 g 192 t 2 others  
ORIGIN  
Query Match 20.0%; Score 305.4; DB 14; Length 907;  
Best Local Similarity 59.8%; Pred. No. 9.2e-69;  
Matches 530; Conservative 0; Mismatches 356; Indels 1; Gaps 1;

QY 217 AAGCCGATACCCATGACGAGATTTTGAAGGTTTAACTTTAATTTGACCGAATCCAG 276  
Db 15 AGGCTGACACTCAGATGAAATCTCGAGGGCTGAATTTCAACCTCAGCGAGATTCGG 74  
QY 277 AAGCCCAATTCACGAGGGTTTCAAGAGTTGTTGAGAATTTGAATCAACCTGATTC 336  
Db 75 AGGCTCAGATCCATGAAGGTTCCAGGAACCTCTCGCTCAACCCAGCCAGACGCC 134  
QY 337 AATTGCAATTAATCTACTGTTAAGGTTTATTTTGTCTGAAGGTTTAAATTTGGTTGACA 396  
Db 135 AGCTCCAGCTGACCCGCGCAATGGCTGTCTCTCAGCGAGGGCTTGAAGCTAGTGATA 194  
QY 397 AATTCTCTAGAACGCTCAAGAACTATATCATAGTCAGGCTTTTACCGTTAATTTGGTG 456  
Db 195 AGTTTTTGGAGGATGTTAAAAGTTTGTACACTCAGAACCTTCACTGTCAACTTCGGGG 254  
QY 457 ATACTGAGGAAGCTAAAAAGCAAATTAATGATTATCTTCGAAAGCCACCCAGGCTAAACA 516  
Db 255 ACACCGAAGAGGCCAAGAACAGATCAACGATTACGTGGAGAGGGTACTCAAGGGGAAA 314  
QY 517 TCCTTGACCTAGTTAAAGAATTAGATCGTGATACCGCTCTTCGCACCTAGTTAACTATATTT 576  
Db 315 TTGTGATTTGGTCAAGGAGCTTGACAGAGACACAGTTTTCCTCTCGGTGAATTACATCT 374

QY 577 TTTTCAAGGTAAGTGGGAAGCTCTTTTCGAGGTAAAGATACACTGAAGAGAAAGATTTTC 636  
 Db 375 TCTTTAAAGCAAAATGGGAGAGACCTTTGAAGTCAAGGACACCGAGGAAGAGACTTCC 434  
 QY 637 ATGTTGATCAAGTACTACTGCAAGTTCCTCAATGATGAAGAAAGTGGGTATGTTCAATA 696  
 Db 435 ACGTGGACAGGTGACCAACCGTGAAGGTGCGCTATGATGAAGCGTTTATAGGATGTTTAAACA 494  
 QY 697 TTTACATTTGCAAAAAATTAAGTCTTGGGTCCTTAAATGAAGTATTTAGGTAACGCTA 756  
 Db 495 TCCAGCATGTGAAGAGCTGTCCAGCTGGGGTCTGCTGATGAATATCTGGGCATGCCA 554  
 QY 757 CTGCTATTTTATTTTATTTTACAGACGAAGTAAAGTCAAGTCTTCAACATTTTAGAATGAGTTGACTC 816  
 Db 555 CGCCATCTTCTCTGCTGCTGATGAGGGAACCTACAGCACCTGGAAATGAATCAACCC 614  
 QY 817 ATGACATTTATTAATTTTATGAGAACGAGGATCGTGTAGGCGTCTCTGCACTGC 876  
 Db 615 ACGATATCATCAACCAAGTTCCTGGAAATGAAGACAGAAAGTCTGCCAGCTTACATTTAC 674  
 QY 877 CAAAGTTAAGTATCACCGGTACTTACGACTTAAATCTGTTTAAAGGCCAGTTAGGTATTA 936  
 Db 675 CAAAACTGTCCATCTGGAACCTATGATCTGAAGAGGCTCCTGGTCAACTGGGCATCA 734  
 QY 937 CCAAAAGTTTTTCTAACCGTCCGATTTGAGTGGTGTACTGAAGAAAGCTCCATTTAAAT 996  
 Db 735 CTAAGTCTTCAGCAATGGGCTGACCTCTCCGGGTACACAGAGGAGGCCCTGAAGC 794  
 QY 997 TGAGTAAAGCTGTTCACAAAGCGGTCTTAAGTATGATGAAGAGGTTACCGAGGCCCGC 1056  
 Db 795 TCTCAGGAGGC-GTGCAATAGGCTGTGCTGACCATCGACGAGAAAGGACTGAAGCTGCTG 853  
 QY 1057 GGGCTATGTTCTCGAAGCTATTCCTCAATGAGCATTCACACAGAAAGTT 1103  
 Db 854 GGGCAATGTTTGAAGGCAATACCATGCTCAATCCCCCGAGGT 900

RESULT 7  
 B0653587  
 LOCUS AGENCOURT\_8349591 NIH\_MGC\_100 Homo sapiens cDNA linear EST 15-JUL-2002  
 DEFINITION 5', mRNA sequence.  
 ACCESSION B0653587  
 VERSION B0653587.1 GI:21777759  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 895)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: CGAP (Stanford)  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LCM2482 row: f column: 07  
 High quality sequence stop: 708.  
 Location/Qualifiers  
 1. 895  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6284550"  
 /clone\_lib="NIH\_MGC\_100"  
 /tissue\_type="hepatocellular carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: liver; Vector: pOTB7; Site\_1: XhoI; Site\_2: "

## FEATURES

source

RESULT 8

ECORI: cDNA made by oligo-dT priming. Directionally cloned into EORI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT 227 a 248 c 230 g 190 t  
 ORIGIN

Query Match 18.8%; Score 286.8; DB 14; Length 895;  
 Best Local Similarity 60.1%; Pred. No. 6.6e-64;  
 Matches 477; Conservative 0; Mismatches 317; Indels 0; Gaps 0;

QY 12 GAAGACCCCTCAAGGCGACCGCGCTCAAAAAACCGACACCAAGTCATCACACCAAGACCAT 71  
 Db 95 GAGGATCCCGAGGAGATGCTGCCAGAGACAGATACATCCACCACCATGATCAGGATCAC 154  
 QY 72 CGGACTTTTAATAAAATTAATCTCCAAATTTAGCCGAATTTGCTTTTCTTTTGTATAGACAA 131  
 Db 155 CCAACCTTCAACAAGATCACCCCAACCTGGCTGAGTTCCGCTTACGCTATATACGCCAG 214  
 QY 132 TTAGCTCATCAAGATTAATCTACTTAACATTTTITTAGTCTCTGTTTCTATTGCCACTGCT 191  
 Db 215 CTGGCACACCGTCCCAACAGCACCAATATCTTCTCTCCCAAGTACGATCGCTACAGCC 274  
 QY 192 TTCGCCATGTTGAGTTAGTACTAAAGCGGATACCCATGAGGAGATTTTAAAGAGTTTA 251  
 Db 275 TTTGCAATGCTCTCCCTGGGACCAAGGCTGACACTCAGATGAATCTCTGGAGGCTG 334  
 QY 252 AACTTTAATTTGACCAGAAATCCAGAAAGCCCAAAATTCAGAGGGTTTTCAGAGTTGTTG 311  
 Db 335 AATTTCAACCTCACGGAGATTCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACCTCTC 394  
 QY 312 AGAATCTTGAATCAACCTGATCTCAATTCGAATTAACCTACTGTTAAACGTTTATTTTG 371  
 Db 395 CGTACCTTCAACGACGACAGCCAGCTCCAGCTGACCAACCGGCANTGGCTGTCTCTC 454  
 QY 372 TCTGAAGGTTTAAATTTGGTTGACAAATTCCTAGAAGAGCTCAAGAAACTATATATAGT 431  
 Db 455 AGCGAGGCTTGAAGCTAGTGGATAAGTTTGGAGGATGTTTAAAGTTGTACCACTCA 514  
 QY 432 GAGGCTTTTACCGTTAAATTTTGGTGTACTGAGGAAGCTTAAAGCAAAATTAATGATTAT 491  
 Db 515 GAAGCCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 574  
 QY 492 GTTGAGAAAGGCCACCGAGGTAAGATCGTTGACCTAGTTTAAAGAAATAGATCGTGATAC 551  
 Db 575 GTGGAAGAGGTAAGTCAAGGGAATTTGGATTTGGTCAAGGAGCTTGACAGAGACACA 634  
 QY 552 GTCTTCGCACTAGTTAACTATATTTTTCAGGGTAAGTGGGAAGCTCCTTTTCGAGGTT 611  
 Db 635 GTTTTGTCTCTGTGAATTACATCTTCTTAAAGGCAATTTGGAGAGAGCCCTTTGAAGTC 694  
 QY 612 AAAGATACGAAGAGGAGAGATTTTCATGTTGATCAAGTTACTACTCTCAAGTTTCAAGT 671  
 Db 695 AAGGACACCGAGGAAGAGGACTTCCACGTGGACAGGTGACCAACCGGAGGTCCTATG 754  
 QY 672 ATGAAAGAGCTGGGTATGTTCAATATTTCAACATTTGCAAAATTAAGTTCTTTGGGTTTA 731  
 Db 755 ATGGAGCGTTTAGCATGTTTAACTCCAGCACTGTGAAGAGCTGTCCAGCTGGGTGCTG 814  
 QY 732 TTAATGAAGTATTTAGGTAAAGCTACTGCTATTTTATTTTACCACAGAGGTAAGCTT 791  
 Db 815 CTGATGAATACCTGGGGCATGCCACCGCATCTTCTTCTGCTGGATGATGAGGAAACTA 874  
 QY 792 CAACATTTAGAGAA 805  
 Db 875 CAGCACCTGGGAAA 888



B0646142  
 LOCUS B0646142 982 bp mRNA linear EST 15-JUL-2002  
 DEFINITION AGENCOURT\_8492569 NIH\_MGC\_100 Homo sapiens cDNA clone IMAGE:6296341  
 5' mRNA sequence.  
 ACCESSION B0646142  
 VERSION B0646142.1 GI:21770314  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 982)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue procurement: CGAP (Stanford)  
 cDNA library preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LNCM2504 row: a column: 14  
 High quality sequence stop: 647.  
 Location/Qualifiers  
 1..982  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NIH\_MGC\_100"  
 /tissue\_type="hepatocellular carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: liver; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Size-selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the laboratory  
 of Gerald M. Rubin (University of California, Berkeley)  
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 Library."  
 BASE COUNT 267 a 259 c 238 g 217 t 1 others  
 ORIGIN  
 Query Match 18.7%; Score 285.2; DB 14; Length 982;  
 Best Local Similarity 59.8%; Pred. No. 1.8e-63;  
 Matches 496; Conservative 0; Mismatches 333; Indels 1; Gaps 1;  
 QY 303 GAGTTGTTGAGAACTTGAATCAACCTGATTTCTCAATTGCAATTAACTACTGTGTAACGGT 362  
 DB 1 GAACCTCTCCGTACCTCAACCCAGCAGACAGCCAGCTCCAGCTGACCCGCGCAATGGC 60  
 QY 363 TTATTTTGTCTGAAGGTTTAAATTTGTTGACAAATTCCTAGAACGCTCAAGAACTA 422  
 DB 61 CTGTTCTCACGAGGGGCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAGATTG 120  
 QY 423 TATCATAGTAGAGGCTTTTACCGTTTAAATTTGGTGATCTACTGAGGAAGCTAAAGCAAT 482  
 DB 121 TACCACCTAGAAGCTTCACTGTCACTCGGGACACCGAAGGCGCAAGAACATC 180  
 QY 483 AATGATTATGTTGAAGAGGCCACCGGTAAGATCGTTGACCTAGTTAAAGAAATTAGAT 542  
 DB 181 AACGATTACGTGGAGAGGGGTACTCAAGGGAATAATTTGGATTGGTCAAGGAGCTTGAC 240  
 QY 543 CGTGATACCGCTTTCGCACTAGTTAACTATATTTTTTCAAGGTAAGTGGGACGTCCT 602  
 DB 241 AGACACACAGTTTTTGTCTGGTGAATTACATCTCTTTAAAGGCAATGGGAGAGACCC 300  
 QY 603 TTCGAGGTTAAAGACTACTGAAGAGGAGATTTTCATCTGTGATCAAGTTACTACTGTCAA 662  
 DB 301 TTTCAAGTCAAGGACACCGGAGGAGAGGACTTCCACGTGGACCCAGGTGACCCCGTGAAG 360

QY 663 GTTCCAATGATGAAAAGACTGGGTATGTTCAATATTCAACATTTGCAAAAAATTAAGTTCT 722  
 DB 361 GTGCTATGATGAAGCGGTTTAGGCATGTTTAACTCCAGCACTGTAAGAAGCTGTCAGC 420  
 QY 723 TGGGTCATTATTAAAGAGTATTTAGGTAAACGCTACTGCTATTTTTTTTACCAAGCAA 782  
 DB 421 TGGGTGCTGCTGATGAATACTCGGCAATGCCACCCCATCTTCTTCTCCCTGATGAG 480  
 QY 783 GGTAAAGCTTCAACATTTAGAGAATGAGTTGACTCATGACATATTACTAAATTTTATAG 842  
 DB 481 GGGAACTACAGCACTTGGAAATGAATCACTCCACCATATCATCAACCAAGTTCTCTGAA 540  
 QY 843 AACGAGGATCTCTAGCGCTTCTGCACTGCCAAAGTTAAGTATCACCGGTACTTAC 902  
 DB 541 AATGAAGACAGAGGCTCTGCCAGCTTACATTTACCAAACTGTCCATTACTGGAACCTAT 600  
 QY 903 GACTTAAATCTGTTTTAGGCCAGTTAGGTATTACCAAGT-TTTTTCTAACGGTCCGA 961  
 DB 601 GATCTGAAGAGCGTCTGGGTCACTGGGATCACTAAGGTCTTTTCAGCAATGGGCTGA 660  
 QY 962 TTTGAGTGGTGTACTGAAGAAGCTCCATTAAATTTAGTAAAGCTGTTCCAAAAGCGCT 1021  
 DB 661 CCTCTCGGGGTACAGAGGAAGCACCCCTGAAGCTCTCCAGGCGGTGATAGGCTGT 720  
 QY 1022 CTTAACTATTGATGAAAAGGTACCGAGGCGCGGGCGGTATGTTCTCGGAAGCTATTC 1081  
 DB 721 GCTGACCATCGACGAGAAAGGAGCTGAAGCTGCTGGGCCATGTTTGTAGAGCCATACC 780  
 QY 1082 AATGACATTCCACCAAGTTAAATTTAATAAACCATTCGTTTTTCTGA 1131  
 DB 781 CATGCTCTATCCCTCCCGGAGGCGAGTTCACCAAAACCTTTGTGTTCTTTA 830

RESULT 9  
 BM924019  
 LOCUS BM924019 985 bp mRNA linear EST 12-MAR-2002  
 DEFINITION AGENCOURT\_6709923 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5760183  
 5' mRNA sequence.  
 ACCESSION BM924019  
 VERSION BM924019.1 GI:19374398  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 985)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue procurement: Life Technologies, Inc.  
 cDNA library preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLM12806 row: m column: 16  
 High quality sequence stop: 707.  
 Location/Qualifiers  
 1..985  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5760183"  
 /clone\_lib="NIH\_MGC\_116"  
 /lab\_host="DH10B"  
 /note="organ: pooled colon, kidney, stomach; Vector:  
 pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
 source anonymous pool of 3 colons, age 26 yo male, 49 yo  
 female, 71 yo male colon; 46 yo male kidney, and pool of 2  
 stomachs, 62 yo male and 70 yo female. Library is

FEATURES  
 source

oligo-dt primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH\_MGC Library.

BASE COUNT 258 a 270 c 242 g 213 t 2 others  
ORIGIN

Query Match 18.6%; Score 284.2; DB 14; Length 985;  
Best Local Similarity 60.3%; Pred. No. 3.2e-63;  
Matches 503; Conservative 0; Mismatches 329; Indels 2; Gaps 2;

QY 12 GAAGACCTTCAAGGCGAGCGCTCAAAAAACCGACAGCTATCATCAGCAAGACCAT 71  
DB 92 GAGGATCCCGAGGAGATGCTGCCAGAGAGACAGATATCCACCATGATCAGGATCAC 151  
QY 72 CGGACCTTTAATAAATTTACTCCAAATTTAGCCGAATTTGCTTTTCTTGTATAGACAA 131  
DB 152 CCAACCTTTCAACAAGATCACCCCAACCTGCTGAGTTCGCTTCAGCCTATACCGCCAG 211  
QY 132 TTAGCTCATCAAAAGTAATTTCTACTAACATTTTTTTTAGTCTGTTTCTATTTGCCACTGCT 191  
DB 212 CTGGCAGCACAGTCCCAACAGCAGCAATATCTTCTCCCGAGTGAGCATCGCTACAGCC 271  
QY 192 TTGGCCATGTTGAGTTAGTACTAAAGCCGATACCCATGACGAGATTTTGAAGGTTTA 251  
DB 272 TTTGCAATGCTCTCCCTGGGACCAAGGCTGACACTCAGATGAATCCTGGAGGCGCTG 331  
QY 252 AACTTTAATTTTCAACGAAATCCAGAACGCCAAATTTACAGAGGTTTTTCAAGAGTTGTTG 311  
DB 332 AATTTCAACCTCAGGAGATTTCCGAGGCTCAGATCCATGAAGGCTTCCAGGAATCCTC 391  
QY 312 AGAATCTTGAATCAACCTGATTTCTCAATTTGCAATTAATTAAGTACTGTTAAGGTTTATTTG 371  
DB 392 CGTACCTTCAACAGCAGCAGCAGCTCCAGCTGACCGCGCAATGGCGCTGTCTC 451  
QY 372 TCTGAGGTTTAAATTTGTTGACAAATTTCTVAGAGAGCTCAAGAACTATATCATAGT 431  
DB 452 AGCGAGGCGCTGAGCTAGTGTATGATGATTTTGGAGGATGTAAAGATTTGTACCACTCA 511  
QY 432 GAGGCTTTTACCGTTAATTTTGGTGATGATGAGGAGCTTAAAGAACTAATTAATGATTAT 491  
DB 512 GAAGCCTTCTCAGTCACTTCCGAGCAGCGAGAGGCGCAAGAAACAGATCAACGATTAC 571  
QY 492 GTTGAGAAAGGACCCAGGTAAGTCTGTGACCTAGTTTAAAGATTTAGATCGTGATACC 551  
DB 572 GTGGAGAGGGTACTCAAGGGAATTTGGGATTTGGTCAAGGAGCTTGACAGACACA 631  
QY 552 GTCTTCGCACTAGTTAACTATATTTTTTTTCAAGGGTAAAGTGGGAACTGCTTTTCGAGGTT 611  
DB 632 GTTTTTGCTCTGTGATTTACATCTTCTTTTAAAGGCAATGGAGAGACCCCTTTGAAGTC 691  
QY 612 AAAGATCTGAAGAGGAAGATTTTCAATTTGATGATCAAGTTACTACTGTTCAAGATTTCAATG 671  
DB 692 AAGGACACCCAGGAGAGGACTTCCACGTGGACAGCTGACCCCGTGAAGGTGGCTATG 751  
QY 672 ATCAAGAGCTGGGTATGTTCAATATTTCAACATTTGCAAAATTAAGTTTCTTGGGTCTTTA 731  
DB 752 ATGAAGGTTTGAAGATTTTAACTTCAATCTCAGCACTGTAAAGCTGTTCCAGCTGGGTGCTG 811  
QY 732 TTAATGAAGTATTTAGGTAACTGCTGCTAT-TTTTTTTTTACAGAGCAAGGTAAGCT 790  
DB 812 CTGATGAATACCTGGCAATGCCACGGCATCTTCTTCTGCTGATGAGGGGAACT 871  
QY 791 TCAACATTTAGAGATGAGTTGACTCATGACATTTACTAAATTTTATAGAGAA 844  
DB 872 ACAGGCACTGGAAATGAACCTCAC-CACGATATCATCAAGGTTTCTGCTGAA 924

RESULT 10  
BQ958958 924 bp mRNA 1linear EST 21-AUG-2002  
LOCUS

DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AGENCOURT\_10034701 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6483305  
5', mRNA sequence.  
BQ958958  
BQ958958.1 GI:22374436  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 924)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: DCTD/DP  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1CM2667 row: k column: 18  
High quality sequence stop: 586.  
Location/Qualifiers

1..924  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6483305"  
/clone\_lib="NIH\_MGC\_40"  
/tissue\_type="carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: prostate; Vector: pOTB7; Site:1: XhoI;  
Site:2: EcoRI; cDNA made by oligo-dr priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

BASE COUNT 251 a 253 c 226 g 193 t 1 others  
ORIGIN

FEATURES  
source

Query Match 18.6%; Score 283.8; DB 14; Length 924;  
Best Local Similarity 58.4%; Pred. No. 4e-63;  
Matches 513; Conservative 0; Mismatches 363; Indels 2; Gaps 1;

QY 84 AAAATTACTCCAAATTTAGCGGAATTTGCTTTTCTTTGTATAGACAAATAGCTCAATCAA 143  
DB 2 AAGATCACCCCAACCTGGCTGAGTTCGCTTACGCTATATACCGCCAGCTGGCACACCAG 61  
QY 144 AGTAATTTCTACTAACAATTTTTTTTAGTCCTGTTTCTATTGCCACTGCTTTGCCCATGTTG 203  
DB 62 TCCAACAGCAGCAATATCTTCTTCTCCCGAGTGCATCGCTACAGCTTTGCAATGCTC 121  
QY 204 AGTTTAGTACTAAAGCGGATACCCATGACGAGATTTTAGAAGGTTTAAACTTTTAAATTTG 263  
DB 122 TCCCTGGGGACCAAGGCTGACACTCAGATGAATCTCGAGGCGCTGAATTTCAACCTC 181  
QY 264 ACCGAAATCCCAAGAGCCCAAAATTCACGAGGGTTTTCAAGAGTTGTTGAGAACTTTGAAT 323  
DB 182 ACGGAGATTCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACCTCCCGTACCCCTCAAC 241  
QY 324 CAACCTGATTTCTCAATTCGAATTAAGTACTGTAACGCTTTATTTTGTCTGAAGGTTTA 383  
DB 242 CAGCCAGACAGCAGCTCCAGCTGACCCAGCGCAATGGCCTTCTCCTCAGCAGGGGCTG 301  
QY 384 AAATTTGTTGACAAATTCCTAGAGACGCTCAAGAACTATATCATAGTGGGCTTTTACC 443  
DB 302 AAGCTAGTGGATAGTTTTTTTGGAGATGTTTAAAGATTTGTACCACCTCAGAGGCTTCACT 361  
QY 444 GTTAATTTTGGTGATCTAGGAGAGCTAAAAGCAAAATTAATGATTATGTTGAGAAAGGC 503

Db 362 GTCAACTTCGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTACGTTGAGAGGCT 421

Qy 504 ACCAGGGTAAGATCGTTGACCTAGTAAAGAAATAGATCGTATACCGCTTCGCACTA 563

Db 422 ACTCAAGGGAAATGTGGATTGGTCAAGAGGCTTTGACAGAGACACAGTTTTTGTCTG 481

Qy 564 GTTAACTATATTTTTTCAAGGTAAGTGGGAACGTCCTTTCAGGTTAAAGATCTGAA 623

Db 482 GTCAATTACATCTCTTTAAAGCAANTGGAGAGACCCCTTGAAGTCAAGACACCGAG 541

Qy 624 GAGGAAGATTTTCATGTTGATCAAGTTACTGTCTCAAAAGTTCCAAATGATGAAGAACTG 683

Db 542 GAAGAGGACTTCCACGTGGACGAGGCGACCGCTGAAGTGCTATGATGAAGGCTTTA 601

Qy 684 GGTATGTTCAATATCAACATTTGCAAAAATTAAGTTCTTGGGCTCTTATTAATGAAGTAT 743

Db 602 GGATGTTTAAACATCCAGCACCTGTAAGAAGCTGTCCAGCTGGGTGCTGTGATGAATAC 661

Qy 744 TTAGGTAAGGCTACTGCTATTTTTTTTTTACCAGAGCAAGGTAAGCTTCCAAACATTTAGAG 803

Db 662 CTGGCAANTGCCACCGCATCTTCTCTGCTGATGAGGGNAACACTACAGCACCTGGAA 721

Qy 804 AATGAGTGTACTCATGACATTAATTAATAATT--TTTAGAGAACGAGGATCGTCGTAGG 861

Db 722 AATGAAGTCAACCCAGCATATCATCACNCAAGTTCCCTGGAAATGAAGACAGAAAGTCCTG 781

Qy 862 CTCTCTGACCTGCCAAAGTTAAGTATCACCGGTACTTTACGACITTAATAATCTGTTTAG 921

Db 782 CCACCTTAGATTTACCCAAACTGTCCATTACTGGAAACTATGATCTGAAGAGCGCTCCTG 841

Qy 922 GCCAGTTAGGTATACCAAAAGTTTTTCTTAACGGTGCC 959

Db 842 GTCAACTGGGCATCACTAAGGCTTCAGCAATGGGGC 879

RESULT 11  
BM924813  
LOCUS  
DEFINITION  
AGENCY: 6653921 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5761267  
5' mRNA sequence.

ACCESSION  
BM924813  
VERSION  
BM924813.1 GI:19375192  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1194)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
DNA distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM12809 row: j column: 20  
High quality sequence stop: 662.  
Location/Qualifiers  
1..1194  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5761267"  
/clone\_lib="NIH\_MGC\_116"  
/lab\_host="DH10B"  
/note="Organ: pooled colon, kidney, stomach; Vector: PCMV-SPORT0; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2

## FEATURES

Source

stomachs, 62 yo male and 70 yo female. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by c. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH\_MGC Library."

BASE COUNT 321 a 303 c 328 g 242 t  
ORIGIN

Query Match 18.0%; Score 274.8; DB 14; Length 1194;  
Best Local Similarity 61.2%; Pred. No. 9.7e-61;  
Matches 444; Conservative 0; Mismatches 282; Indels 0; Gaps 0;

Qy 448 ATTTTGGTACTGAGGAAGCTAAAAAGCAAATTAATGATTGTTGAGAAAGGCACCC 507

Db 7 AATTTCCGGATCCGAGAGGCCAAGAAACAGATCAACGATTACGTGGAGAGGGTACTC 66

Qy 508 AGGTAAGATCGTTGACCTAGTTAAAGNATTTAGATCGTGATACCGTCTTCCCATAGTTA 567

Db 67 AAGGAAAATTTGTGGATTGTGTCAGGAGCTTGACAGAGACACAGTTTTTGTCTGTGTA 126

Qy 568 ACTATATTTTTTCAAGGTTAAGTGGGAACGTCCTTTTCGAGGTTAAAGATACTCAAGAGG 627

Db 127 ATTACATCTCTTTAAAGGCAATGGGAGAGACCCCTTTGAAGTCAAGACACCCAGGAAG 186

Qy 628 AAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTTCCAAATGATGAAAGAGCTGGGTA 687

Db 187 AGGACTTCCACGTGACCCAGGTGACCCCGTGAAGGTGCTTATGATGAAGCGTTTAGGCA 246

Qy 688 TGTTCAATATTTCAACATTTGCAAAAATTAAGTTCTTGGGCTTTATTAATGAAGTATTAG 747

Db 247 TGTTTACATCCAGCACCTGTAAAGAAAGCTGTCCAGCTGGGTGCTGCTGATGAATACCTGG 306

Qy 748 GTAAGCTACTGCTATTTTTTTTTTACCAGAGCAAGGTAAAGCTTCAACATTTTAGAGAAATG 807

Db 307 GCATGGCACCGCCATCTTCTCTGCTGATGAGGGGAAACTACAGCACTGGAATG 366

Qy 808 AGTTGACTCATGACATTAATTAATAATTTTATAGAAACGAGGATCGTCGTAGCGCTTCTC 867

Db 367 AACTCACCCAGCATATCATCAACCAAGTTCTTGGAAATTAAGACAGAAAGGTCTGCCAGCT 426

Qy 868 TGCACCTGCCAAAGTTAAGTATCACCGTACTTACGACTTAAATCTGTTTAGGCCAGT 927

Db 427 TACATTTACCCAAAGTGTCCATTACTGGAACCTATGATCTGAAGAGCGTCTCTGGGTCAAC 486

Qy 928 TAGTATTACCAAAAGTTTTTTTAAACGGTGGCGGATTTGAGTGGTGTACTGAAGAAGCTC 987

Db 487 TGGGCATCACTAAGTCTTCAGCAATGGGCTGACCTCTCCGGGTGTCACAGAGGAGGCAC 546

Qy 988 CATTAATAATTGAGTAAAGCTGTTTCAAAAGCGCTTTAACTATTGATGAAGAGGGTACCG 1047

Db 547 CCCTGAAGCTCTCCAAGGCGGTGCATTAAGGCTGTGCTGACATCGACGAGAGAGGACTG 606

Qy 1048 AGGCGCGCGGCGCTGCTTCTGGAAGCTATTCCAATGAGCATTTCCACCAAGAGTTAAAT 1107

Db 607 AAGTGTCTGGGGCCATGTTTTTAGAGGCCATPACCCATGTCTATCCCCCGGAGGTCAAGT 666

Qy 1108 TTAATAAACCATTCGTTTTTTTCTGATGATCGAGCAACACATTAAGAGCCCATTTGTTATGG 1167

Db 667 TCAACAAACCTTTGTCTTCTTAATGATTGAACAAATACCAAGTCTCCCCCTCTTCATGG 726

Qy 1168 GTAAGG 1173

Db 727 GAAAA 732

## RESULT 12

BQ650189

LOCUS

DEFINITION

AGENCY: 8298326 NIH\_MGC\_100 Homo sapiens cDNA clone IMAGE:6269613

5' mRNA sequence.

ACCESSION

BQ650189

907 bp mRNA linear EST 15-JUL-2002  
AGENCY: 8298326 NIH\_MGC\_100 Homo sapiens cDNA clone IMAGE:6269613  
5' mRNA sequence.

```
VERSION B0650189.1 GI:21774361
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL 1 (bases 1 to 907)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: CGAP (Stanford)
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2443 row: g column: 22
High quality sequence stop: 650.
FEATURES
source
Location/Qualifiers
1..907
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_100"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pOTB7; Site:1; XhoI; Site:2;
EcORI; CDNA made by oligo-dT priming. Directionally cloned
into EcORI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 245 a 237 c 226 g 198 t 1 others
Query Match 17.9%; Score 273.4; DB 14; Length 907;
Best Local Similarity 59.4%; Pred. No. 2.1e-60;
Matches 498; Conservative 0; Mismatches 337; Indels 3; Gaps 2;
QY 336 CAATTGCAATTAAGTAACTACTGCTAACTGCTTTATTTTGTCTGAGGTTTAAAAATGTTGAC 395
DB 7 GAGCTCCAGCTGACACCGGCAATGGCTGTTCTCAGCGAGGGCTGAAGCTAGTGGAT 66
QY 396 AAATCTCTAGAGAGCTCAAGAACTATATCATAGTAGGCTTTTACCGTTAAATTTGGT 455
DB 67 AAGTTTTTGGAGGATGTTAAAAAGTTGTACCACCTCAGAAAGCTTCACTGTCAACTTCGGG 126
QY 456 GATAGTGAAGAGCTAAAAGCAATTAATGATTATTTGAGAAAGGACCCAGGTTAAG 515
DB 127 GACACGGAAGGCCAAGAAACAGATCAACGATTACGTTGAGAGAGGTTACTACAGGGAAA 186
QY 516 ATCGTTGACCTAGTTAAAGAAATAGATCGTATACCGTTTCGCGCTAGTTAACTATATT 575
DB 187 ATGTGGAATTTGGTCAGAGAGCTTGACAGAGACACAGTTTTTGTCTGTGTAATTACATC 246
QY 576 TTTTCAAGGGTAAGTGGGAAGCTGCTTTTCGAGGTTAAAGATACATGAAGAGGAAGATTTT 635
DB 247 TTTCTTTAAAGGCAATGGGAGAGACCTTTTGAAGTCAAGGACACCGGAGAGAGGACTTC 306
QY 636 CATGTTGATCAAGTTACTACTGTCAAAGTTCCAAATGATGAAAGACTGGGTATGTTCAAT 695
DB 307 CAGGTGGACAGGTGACCGGTAAGGTCCTATGATGAACCGTTAGGATGTTTAC 366
QY 696 ATTCACATTTGCAAAAAATTAAGTTCTTGGGCTTTATTAATGAAGTATTTAGGTAAACGCT 755
DB 367 ATCCAGCACTGTAAAGAGCTGTCCAGCTGGTGTCTGATGAATACCTGGGCAATGCC 426
```

```
QY 756 ACTGCTATTTTTTTTACCAGCAAGGTAAAGCTTCAACATTTAGAGAATGAGTTGACT 815
DB 427 ACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTACAGCACCTGGAAATGAACCTCAC 486
QY 816 CATGACATATTACTAAATTTTTTAGAGAACGAGGATCGTCTGAGCGTTCTCTGCACCTG 875
DB 487 CACGATATCATCAACCAAGTTCTCTGAAAAATGAAGACAGAGGTCTGCCAGCTTACATTTA 546
QY 876 CCAAAGTTAAGTATACACCGGTACTTACGACTTAAATCTGTTTAGGCCAGTTAGGTATT 935
DB 547 CCCAAACTGTCATTTACTGGAACCTATGATCTGAAGACGCTCTGGGTCAACTGGGCATC 606
QY 936 ACCAAGTTTTTCTTAACGGTCCGATTTGAGTGTGCTGTACTGAAGAAAGCTCCATTTAAA 995
DB 607 ACTAAGGCTTTCAGCAATGGGGCTGACCTCTCCGGNGTCACAGAGAGGACCCCTGAAG 666
QY 996 TTGAGTAAAGCTGTTTACAAAGCCGCTTCTTAATGATGATAAGAGGTTACGAGGCCGCC 1055
DB 667 CTCTCCAAGCGCTGCATTAAGGCTGTGTGACCATCGACGAGAAAGG-ACTGAAACTGCT 725
QY 1056 GGGCTATGTTCTTGAAGCTATTCATGAGCATTCACACAGAGTTTAAATTTAA--TA 1113
DB 726 GGAGCCATGTTTTTAGAGGGCTTACCCATGCTATTCCTCCCGGAGTCAAGTTTACCAC 785
QY 1114 AACCATTCGTTTTTCTGATGATCGAGCAGACACATAAAAGCCCATTTGTTTATGGGTAA 1171
DB 786 ACCCTTGTCTCTTTAATGATGGACCAATACCAAGTCTCCCTCTTCCGGGGAA 843
```

```
RESULT 13
B0648524
LOCUS B0648524 959 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8297828 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6269919
5', mRNA sequence.
ACCESSION B0648524
VERSION B0648524.1 GI:21772696
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: CGAP (Stanford)
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2444 row: d column: 16
High quality sequence stop: 674.
```

```
FEATURES
source
Location/Qualifiers
1..959
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6269919"
/clone_lib="NIH_MGC_100"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pOTB7; Site:1; XhoI; Site:2;
EcORI; CDNA made by oligo-dT priming. Directionally cloned
into EcORI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life technologies). Note: this is a NIH_MGC
Library."
```

BASE COUNT ORIGIN	246 a	276 c	245 g	192 t
Query Match	17.7%; Score 270.4; DB 14; Length 959;			
Best Local Similarity	59.3%; Pred. No. 1.3e-59;			
Matches 477; Conservative	0; Mismatches 326; Indels 1; Gaps 1;			
QY	186	ACTGCTTCGCGCATGTTGAGTTAGTACTAAAGCGGATACCATGACGAGATTTTAGAA	245	
DB	1	ACAGCCCTTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCAGATGAATCTGGAG	60	
QY	246	GTTTAACTTTAAATTTGACGAAATCCAGAAAGCCCAATTCACGAGGTTTTCAGAG	305	
DB	61	GGCCTCAATTTCAACCTTCACGGAGATTCGGAGGCTCAGATCATGAAGCTTCAGGAA	120	
QY	306	TTGTTGAGAACTTTGAACTCAACTGATCTCAATTTGCAATTAACACTAGTGGTAACGGTTTA	365	
DB	121	CTCCTCCGTAACCTCAACGAGCAGACCCAGCTCCAGCTGACCAACCGCAATGGCCTG	180	
QY	366	TTTTCTCTGAAGGTTTAAATTTGGTTGACAAATTCCTAGAAGAGCTCAAGAAACTATAT	425	
DB	181	TTCCCTCAGGAGGCTGAGCTAGTGGATTAAGTTTTCGGAGATGTTTAAAGTTGTAC	240	
QY	426	CATAGTGGAGGCTTTTACCCTTAATTTTGGTGATCTAGGAAAGCTAAAAAGCAAAATTAAT	485	
DB	241	CACACAGAGCTTCACTGTCAACTTCGGGGACACCGAGAGAGCCCAAGAAACAGATCAAC	300	
QY	486	GATTATGTTGAGAAAGGACCCAGGTAAGATCGTTGACCTAGTTAAAGAAATGATCGT	545	
DB	301	GATTACGTGGAGAGGGTACTCAAGGAAATTTGGATTTTGGTCAAGGAGCTTGACAGA	360	
QY	546	GATACCGTCTCGCACTACTTAACATATATTTTTCAGGGTAAGTGGGAAGCTCCTTTC	605	
DB	361	GACACAGTTTTCCTGCTGATTAACATCTTCTTTAAAGGCAATGGGAGAGACCCCTT	420	
QY	606	GAGTTTAAAGATGACTGAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGCTCAAGTT	665	
DB	421	GAAGTCAAGGACACCGGAGGAGGAGTCCACGTGGACCAAGTGCACACCGTGAAGGTG	480	
QY	666	CCAATGATCAAAAGACTGGTATGTTCAATATTCACATTTCAACATTTGAAAAATTAAGTTCTTGG	725	
DB	481	CCTATGATGAAGGTTTAGGCTATGTTTAACTCCAGCACTGTAAGAAGCTGCCAGCTGG	540	
QY	726	GTCTTATTAATGAAGTATTTAGTACGCTACTGCTATTTTTCACGAGCAAGGT	785	
DB	541	GTGCTCTGATGAATACCTGGCAATGCCACGCCATCTTCTCTCTGCTGATGAGGG	600	
QY	786	AGCTTCAACATTTAGAGAAATGAGTTGACTCATGACATTAATTAATTTTATGAGAAC	845	
DB	601	AAACTACAGCACTGGAAATGAACCTACCCACGATATCATCAACCAAGTTCTCTGGAAAT	660	
QY	846	GAGGATCGTGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGA-	904	
DB	661	GAAGACAGAGGTCTGCCAGCTTACATTTTACCCTCAACTCTCCATTAAGTGAACCTATGAT	720	
QY	905	CTTAAATCTGTTTATAGGCGAGTTAGGTATTACCAAGTTTTCCTTAACGGTGCCTGATTT	964	
DB	721	CTGAAGAGCGCTCTGGGTCAACTGGGCATCACTAAGGCTTCAGGCCCTTCAGGCTGGACT	780	
QY	965	GAGTGGTGTACTGAAGAGCTCC	988	
DB	781	CCCCGGGGTCAACAGAGGAAGCC	804	
RESULT 14	BQ646948			
LOCUS	BQ646948			
DEFINITION	AGENCOURT_8302495 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6271313			
ACCESSION	BQ646948			
VERSION	BQ646948.1			
KEYWORDS	EST.			
SOURCE	human.			

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 891)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: CGAP (Stanford)

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Plate: LCM2447 row: n column: 18

High quality sequence stop: 672.

Location/Qualifiers

1..891

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:6271313"

/clone\_lib="NIH\_MGC\_100"

/tissue\_type="hepatocellular carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: liver; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT 229 a 247 c 226 g 189 t

Query Match 17.6%; Score 268.6; DB 14; Length 891;

Best Local Similarity 60.3%; Pred. No. 3.7e-59;

Matches 479; Conservative 0; Mismatches 314; Indels 2; Gaps 2;

QY 12 GAAGACCCCTCAAGCGGACGGCTCAAAAACCCAGACCATCATCAGCACCAAGACCAT 71

DB 88 GAGGATCCCGAGGAGATGCTGCCAGAGACAGATACATCCACCATGATCAGATCAC 147

QY 72 CCGACTTTTAAATAAATTTACTCCAAATTTAGCCGAATTTGCTTTTCTTTGTATAGACAA 131

DB 148 CCAACCTTCAACAAGATCACCCCAACCTGGCTGAGTTCGCCCTTACGCCCTATACGCCAG 207

QY 132 TTAGCTCATCAAGTAATTTCTACTAAGATTTTCTTTTAGTCTGCTTTCTATTGCCACTGCT 191

DB 208 CTGGCACACAGTCCACAGACCAATATCTTCTCTCCAGGAGCATCGCTACAGCC 267

QY 192 TTCGCCATGTTGAGTTTAGTACTAAAGCCGATACCCATCAGCAGATTTTAGAAGGTTTA 251

DB 268 TTTGCAATGCTCTCCCTGGGACCAAGCTGACACTCAGCATGAATCTCTGGAGGCGCTG 327

QY 252 AACTTTTAAATTTGACCGAAATCCCAAGAGCCCAATTTACAGAGGTTTTCAGAGTTTGTG 311

DB 328 AATTTCAACCTCACGGAGATTTCCGGAGGCTCAGATCATCAAGGCTTCCAGGAACCTCCTC 387

QY 312 AGAACTTTGAATCAACCTGATTTCTCAATTTGAATTAATTAATTAATTAATTTTGTG 371

DB 388 CGTACCTTCAACAGCAGACAGACAGCTCCAGCTGACCCAGCAATGGCTGTTCTCCTC 447

QY 372 TCTGAAGGTTTAAATTTGGTTGACAAATTTCTAGAGACGCTCAAGAACTATATCATPAT 431

DB 448 AGCAGAGGCGCTGAAGCTAGTGGATAAGTTTTCGAGGATGTTTAAAGTTTGTACCACTCA 507

QY 432 GAGGCTTTTACCGTTTAAATTTTGGTGTACTGAGGAAGCTAAAGCAAAATTAATGATTAT 491

DB 508 GAAGCCTTCTACTGTCAACTTCGGGGACACCCAGAGAGGCCAAGAACAGATCAACGATTAC 567

